

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of

Kazuhiro MOCHINAGA et al.

Serial No. NEW

Filed September 23, 2003

FEATURE QUANTITY EXTRACTING
APPARATUS

: Attn: APPLICATION BRANCH

: Attorney Docket No. 2003_1343A

CLAIM OF PRIORITY UNDER 35 USC 119

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

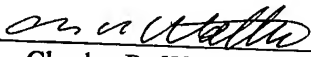
Sir:

Applicants in the above-entitled application hereby claim the dates of priority under the International Convention of Japanese Patent Application No. 2002-277001, filed September 24, 2002, and Japanese Patent Application No. 2003-111121, filed April 16, 2003, as acknowledged in the Declaration of this application.

Certified copies of said Japanese Patent Applications are submitted herewith.

Respectfully submitted,

Kazuhiro MOCHINAGA et al.

By 
Charles R. Watts
Registration No. 33,142
Attorney for Applicants

CRW/asd
Washington, D.C. 20006-1021
Telephone (202) 721-8200
Facsimile (202) 721-8250
September 23, 2003

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日

Date of Application:

2002年 9月24日

出 願 番 号

Application Number:

特願2002-277001

[ST.10/C]:

[JP2002-277001]

出 願 人

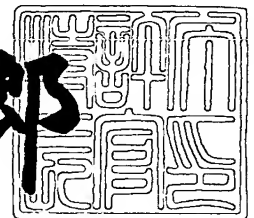
Applicant(s):

松下電器産業株式会社

2003年 6月 2日

特 許 庁 長 官
Commissioner,
Japan Patent Office

太田 信一郎



出証番号 出証特2003-3042681

【書類名】 特許願

【整理番号】 2022540290

【提出日】 平成14年 9月24日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 5/00

【発明者】

【住所又は居所】 大阪府門真市大字門真 1 0 0 6 番地 松下電器産業株式会社内

【氏名】 持永 和寛

【発明者】

【住所又は居所】 大阪府門真市大字門真 1 0 0 6 番地 松下電器産業株式会社内

【氏名】 小川 智輝

【発明者】

【住所又は居所】 大阪府門真市大字門真 1 0 0 6 番地 松下電器産業株式会社内

【氏名】 森 美裕

【特許出願人】

【識別番号】 000005821

【氏名又は名称】 松下電器産業株式会社

【代理人】

【識別番号】 100097445

【弁理士】

【氏名又は名称】 岩橋 文雄

【選任した代理人】

【識別番号】 100103355

【弁理士】

【氏名又は名称】 坂口 智康

【選任した代理人】

【識別番号】 100109667

【弁理士】

【氏名又は名称】 内藤 浩樹

【手数料の表示】

【予納台帳番号】 011305

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9809938

【書類名】 明細書

【発明の名称】 特徴量抽出装置

【特許請求の範囲】

【請求項 1】 音声信号から特徴量を抽出する特徴量抽出装置であって、
前記音声信号から周波数スペクトルを求める周波数変換手段と、
前記周波数変換手段によって求めた前記周波数スペクトルを 1 つ以上の帯域に
分割する帯域分割手段と、
前記帯域分割手段によって分割された帯域の各々から特徴量を抽出する抽出手
段と、
を具備することを特徴とする特徴量抽出装置。

【請求項 2】 前記帯域分割手段は、前記周波数変換手段によって求められた前
記周波数スペクトルを周波数軸上の線形目盛りに従って等間隔に分割することを
特徴とする請求項 1 記載の特徴量抽出装置。

【請求項 3】 前記帯域分割手段は、前記周波数変換手段によって求められた前
記周波数スペクトルを周波数軸上の対数目盛りに従って等間隔に分割することを
特徴とする請求項 1 記載の特徴量抽出装置。

【請求項 4】 前記帯域分割手段は、前記周波数変換手段によって求められた前
記周波数スペクトルを任意の方法にて分割後、ある特定の範囲を含む帯域を選択
することを特徴とする請求項 1 記載の特徴量抽出装置。

【請求項 5】 前記帯域分割手段は、前記周波数変換手段によって求められた前
記周波数スペクトルを帯域の範囲が重複することなく、あるいは接することなく
離散的に分割することを特徴とする請求項 1 記載の特徴量抽出装置。

【請求項 6】 前記抽出手段は、入力される周波数スペクトルからピーク値を検
出し、検出された前記ピーク値をとる周波数を特徴量として出力することを特徴
とする請求項 1 記載の特徴量抽出装置。

【請求項 7】 前記抽出手段は、入力される周波数スペクトルからピーク値を検
出するピーク検出手段と、
前記ピーク検出手段によって検出されたピーク値をとる周波数をある一定時間
遅延させる遅延手段と、

前記遅延手段の出力と前記ピーク値をとる周波数とを比較し、周波数の差分値を特徴量として出力する比較手段と、

を具備することを特徴とする請求項 1 記載の特徴量抽出装置。

【請求項 8】前記抽出手段は、入力される周波数スペクトルからピーク値を検出するピーク検出手段と、

前記ピーク検出手段によって検出されたピーク値をとる周波数をある一定時間遅延させる遅延手段と、

前記遅延手段の出力と前記ピーク値をとる周波数とを比較して得られる周波数の時間軸方向の増減を 2 値の特徴量として出力する比較手段と、

を具備することを特徴とする請求項 1 記載の特徴量抽出装置。

【請求項 9】前記抽出手段は、入力される周波数スペクトルからピーク値を検出するピーク検出手段と、

前記ピーク検出手段によって検出されたピーク値をとる周波数をある一定時間遅延させる遅延手段と、

前記遅延手段の出力と前記ピーク値をとる周波数とを比較し、前記遅延手段の出力を基準として前記ピーク値をとる周波数の変化分がある一定値の範囲を超えるかどうかを 2 値の特徴量として出力する比較手段と、

を具備することを特徴とする請求項 1 記載の特徴量抽出装置。

【請求項 10】前記抽出手段は、入力される周波数スペクトルからピーク値を検出する第 1 のピーク検出手段と、

前記第 1 のピーク検出手段の入力とは異なる帯域における周波数スペクトルからピーク値を検出する第 2 のピーク検出手段と、

前記第 1 のピーク検出手段によって検出された第 1 のピーク値と、前記第 2 のピーク検出手段によって検出された第 2 のピーク値とを比較して得られる前記第 1 のピーク値と前記第 2 のピーク値との差分値を特徴量として出力する比較手段と、

を具備することを特徴とする請求項 1 記載の特徴量抽出装置。

【請求項 11】前記抽出手段は、入力される周波数スペクトルからピーク値を検出する第 1 のピーク検出手段と、

前記第 1 のピーク検出手段の入力とは異なる帯域における周波数スペクトルからピーク値を検出する第 2 のピーク検出手段と、

前記第 1 のピーク検出手段によって検出された第 1 のピーク値と、前記第 2 のピーク検出手段によって検出された第 2 のピーク値とを比較して得られるピーク値の周波数軸方向の増減を 2 値の特徴量として出力する比較手段と、

を具備することを特徴とする請求項 1 記載の特徴量抽出装置。

【請求項 1 2】前記抽出手段は、入力される周波数スペクトルの実効値を計算し、算出された前記実効値を特徴量として出力することを特徴とする請求項 1 記載の特徴量抽出装置。

【請求項 1 3】前記抽出手段は、入力される周波数スペクトルの実効値を計算する実効値計算手段と、

前記実効値計算手段によって計算された実効値を一定時間遅延させるための遅延手段と、

前記遅延手段の出力と前記実効値とを比較し、実効値の差分値を特徴量として出力する比較手段と、

を具備することを特徴とする請求項 1 記載の特徴量抽出装置。

【請求項 1 4】前記抽出手段は、入力される周波数スペクトルの実効値を計算する実効値計算手段と、

前記実効値計算手段によって計算された実効値を一定時間遅延させる遅延手段と、

前記遅延手段の出力と前記実効値とを比較し、実効値の時間軸方向の増減を 2 値の特徴量として出力する比較手段と、

を具備することを特徴とする請求項 1 記載の特徴量抽出装置。

【請求項 1 5】前記抽出手段は、入力される周波数スペクトルの実効値を計算する実効値計算手段と、

前記実効値計算手段によって計算された実効値を一定時間遅延させる遅延手段と、

前記遅延手段の出力と前記実効値とを比較し、前記遅延手段の出力を基準として前記実効値の変化分がある一定値の範囲を超えるかどうかを 2 値の特徴量とし

て出力する比較手段と、

を具備することを特徴とする請求項 1 記載の特徴量抽出装置。

【請求項 1 6】前記抽出手段は、入力される周波数スペクトルを一定時間遅延させるスペクトル遅延手段と、

入力される周波数スペクトルと前記スペクトル遅延手段から出力される周波数スペクトルとの相関値を出力する相関演算手段と、

を具備することを特徴とする請求項 1 記載の特徴量抽出装置。

【請求項 1 7】前記抽出手段は、入力される周波数スペクトルを一定時間遅延させるスペクトル遅延手段と、

入力される周波数スペクトルと前記スペクトル遅延手段から出力される周波数スペクトルとの相関値を計算し、算出した前記相関値の符号を 2 値の特徴量として出力する相関演算手段と、

を具備することを特徴とする請求項 1 記載の特徴量抽出装置。

【請求項 1 8】前記抽出手段は、入力される周波数スペクトルを一定時間遅延させるスペクトル遅延手段と、

入力される周波数スペクトルと前記スペクトル遅延手段から出力される周波数スペクトルとの相関値を出力する相関演算手段と、

前記相関演算手段によって算出される相関値を一定時間遅延させる相関値遅延手段と、

前記相関値遅延手段の出力と前記相関演算手段によって算出される相関値とを比較して得られる相関値の時間軸方向の増減を 2 値の特徴量として出力する比較手段と、

を具備することを特徴とする請求項 1 記載の特徴量抽出装置。

【請求項 1 9】音声信号から特徴量を抽出する特徴量抽出装置であって、
前記音声信号をある一定の時区間ごとに切出す信号切出し手段と、
前記信号切出し手段によって切出された信号を一定時間遅延させる信号遅延手段と、

前記信号遅延手段から出力される信号と前記信号切出し手段によって切出された信号とを入力して特徴量を抽出する抽出手段と、

を具備することを特徴とする特徴量抽出装置。

【請求項 2 0】前記抽出手段は、前記信号遅延手段の出力信号と前記信号切出し手段の出力信号との相関値を計算し出力することを特徴とする請求項 1 9 記載の特徴量抽出装置。

【請求項 2 1】前記抽出手段は、前記信号遅延手段の出力信号と前記信号切出し手段の出力信号との相関値を計算し、算出した前記相関値の符号を 2 値の特徴量として出力することを特徴とする請求項 1 9 記載の特徴量抽出装置。

【請求項 2 2】前記抽出手段は、前記信号遅延手段の出力信号と前記信号切出し手段の出力信号との相関値を計算する相関演算手段と、

前記相関演算手段によって算出された相関値を一定時間遅延させる相関値遅延手段と、

前記相関値遅延手段の出力と前記相関演算手段によって算出された相関値とを比較して得られる相関値の時間軸方向の増減を 2 値の特徴量として出力する比較手段と、

を具備することを特徴とする請求項 1 9 記載の特徴量抽出装置。

【請求項 2 3】音声信号から特徴量を抽出する特徴量抽出装置であって、

前記音声信号から周波数スペクトルを求める周波数変換手段と、

前記周波数変換手段によって得られた前記周波数スペクトルの包絡線を求めるための包絡線検出手段と、

前記包絡線検出手段によって得られた前期周波数スペクトルの包絡線から特徴量を抽出する抽出手段と、

を具備することを特徴とする特徴量抽出装置。

【請求項 2 4】前記抽出手段は、入力される周波数スペクトルの包絡線から極大値あるいは極小値を検出し、検出された極大値あるいは極小値をとる周波数を特徴量として出力することを特徴とする請求項 2 3 記載の特徴量抽出装置。

【請求項 2 5】前記抽出手段は、入力される周波数スペクトルの包絡線から極大値あるいは極小値を検出する極値検出手段と、

前記極値検出手段によって検出された極大値あるいは極小値をとる周波数を入力し、隣接する極値間の間隔を算出し、得られる間隔の値を特徴量として出力す

る間隔算出手段と、

を具備することを特徴とする請求項 2 3 記載の特徴量抽出装置。

【請求項 2 6】前記間隔算出手段によって得られる値に対して、ある基準となる値に対する割合として表現した比を特徴量として抽出することを特徴とする請求項 2 5 記載の特徴量抽出装置。

【請求項 2 7】前記間隔算出手段によって得られる値に対して、前記極値検出手段によって得られた最も低い周波数を基準値として、前記基準値に対する割合として表現した比を特徴量として抽出することを特徴とする請求項 2 6 記載の特徴量抽出装置。

【請求項 2 8】前記間隔算出手段によって得られる値に対して、前記極値検出手段によって得られた最も低い周波数と 2 番目に低い周波数との間隔の値を基準値として、前記基準値に対する割合として表現した比を特徴量として抽出することを特徴とする請求項 2 6 記載の特徴量抽出装置。

【発明の詳細な説明】

【 0 0 0 1 】

【発明の属する技術分野】

本発明は、信号自体が有する固有の特徴量の抽出ならびに特徴量の比較を行う装置およびその方法に関するものであり、詳しくは、音声信号に含まれる特徴量あるいは複数の特徴量の組み合わせた集合を抽出し、他の音声信号と容易に識別することを可能とする装置およびその方法に関するものである。

【 0 0 0 2 】

【従来の技術】

近年、パーソナルコンピュータおよびインターネットの普及とハードディスクドライブなど記憶媒体の大容量・低価格化に伴い、音楽 CD の楽曲をコンピュータの記憶媒体に音楽データとして取り込み、CD-R メディアにコピーしたり、インターネットなどのネットワーク経由で接続されたコンピュータと音楽データの交換や共有をするなどの行為が頻繁に行われている。以上のような行為は楽曲の著作権者に対する権利の侵害であり、したがって、このような行為から著作権を保護するために多くの技術が検討されている。

【0003】

楽曲の著作権保護技術として、現在、一般的に用いられている技術として電子透かし技術がある。電子透かし技術は、音声信号に人間の聴覚ではほとんど知覚できない程度の雑音や歪みとしてデジタル情報を埋め込む技術であり、例えば、DVDオーディオのディスクおよびプレーヤーに既に実用化されている。しかしながら、電子透かしは元の音声信号に対して事前に情報を埋め込んでおく必要があり、また、音声信号自体に直接情報信号を付加するため、元の音声信号を変化させてしまい、例えば、埋め込み情報として、楽曲のタイトル、作曲者、作詞者、歌手名などを埋め込むには非常に膨大な情報量となるため、埋め込み後の音声信号の音質が劣化する可能性がある。

【0004】

以上のような電子透かし技術に対して、近年注目されている技術として音声指紋技術がある。音声指紋技術は、音声信号から人間の指紋のようにその音声信号を特定することができる固有の特徴量（以下音声指紋と呼ぶ）を抽出し、データベースなどにあらかじめ蓄積されている音声指紋と照合して音声信号を識別する技術である。音声指紋技術を用いた例としては、インターネット上でファイル交換を行うソフトウェアにおいて送受信するファイルの内容をチェックしてフィルタリングを行う機能として適用されているものが存在する。

【0005】

ここで、図17および図18を参照しながら、音声指紋技術について説明する。

【0006】

まず、図17は音声信号から抽出された音声指紋と前記音声信号に関する情報を蓄積手段に蓄積する過程を説明するものである。音声指紋技術では、データベースに楽曲のタイトル、作曲者、作詞者、歌手名などの書誌情報や管理情報をあらかじめ保持しておく必要がある。したがって、最初に特徴量抽出装置1701によって音声信号から音声指紋が得られると、続いて、前記音声信号に関する書誌情報や管理情報とともに蓄積手段1702に蓄積される。

【0007】

次に、図 1 8 はある未知の音声信号があるとき、前記未知の音声信号から音声指紋を抽出し、音声信号の特定を行う過程を説明するものである。まず、未知の音声信号は特徴量抽出装置 1 8 0 1 に入力され、前記音声信号の音声指紋が抽出される。抽出された前記音声指紋は、続いて、指紋比較手段 1 8 0 3 に入力される。指紋比較手段 1 8 0 3 では、入力された前記音声指紋と蓄積手段 1 8 0 2 に蓄積されている音声指紋との比較を行う。そして、蓄積されている音声指紋の中から入力された音声指紋と一致するもの、あるいはある基準の範囲内で類似する音声指紋を検出し、検出された音声指紋に関する楽曲情報を出力する。

【 0 0 0 8 】

以上のようにして、音声指紋技術によって音声信号に関する情報を得ることができる。音声指紋技術では、あらかじめ情報を蓄積手段に蓄積する必要があるものの、電子透かし技術と比較して、音声信号自体に情報を埋め込む必要がないため信号を変化させることがなく、また、多くの情報量がある場合にもデータベースに蓄積しておき、容易に取り出すことができる利点がある。

【 0 0 0 9 】

従来、音声指紋として抽出される特徴量には、信号振幅、帯域幅、ピッチ数、メル周波数ケプストラム係数などの物理量が用いられ、各々の平均、標準偏差などの統計的性質を求め、信号の識別を行っている（例えば、特許文献 1 参照）。

【 0 0 1 0 】

【特許文献 1】

米国特許第 5 9 1 8 2 2 3 号

【 0 0 1 1 】

【発明が解決しようとする課題】

しかしながら、音声指紋技術では以下のような課題がある。

【 0 0 1 2 】

音声指紋技術は、音声信号を特定するために音声信号自身から固有の特徴量、すなわち音声指紋を抽出しなければならないが、従来技術で用いられている特徴量は非常に基本的な物理量による実現にとどまっており、音声指紋を実用的なものとするためには、さまざまな信号の加工や外部雑音に対してロバスト性の高い

特徴量を抽出することが課題である。

【0013】

また、音声指紋を実現するに当たって、音声指紋となる特徴量は元の音声信号と比べてできる限り小さいデータ量にて表現することができ、特徴量を抽出する際の演算量を少ないものとするのが課題である。

【0014】

【課題を解決するための手段】

本発明は、音声信号から音声指紋となる特徴量を抽出する特徴量抽出装置である。

【0015】

まず、本発明の特徴量抽出装置は、音声信号から周波数スペクトルを求める周波数変換手段と、前記周波数変換手段によって求めた前記周波数スペクトルを1つ以上の帯域に分割する帯域分割手段と、前記分割手段によって分割された帯域の各々から特徴量を抽出する抽出手段とから構成される。

【0016】

また、前記帯域分割手段は、周波数軸を線形目盛りや対数目盛りに従って等間隔に分割することを特徴とするものであったり、任意の方法にて分割後、特定の帯域のみを選択することを特徴とするものであったり、帯域の範囲が重複したりあるいは接したりすることなく離散的に分割することを特徴とするものである。

【0017】

また、前記抽出手段は、入力される周波数スペクトルからピーク値を検出し、検出されたピーク値をとる周波数を特徴量として抽出することを特徴とするものであったり、入力される周波数スペクトルからピーク値を検出し、検出されたピーク値をとる周波数と遅延手段によってある一定時間遅延させたピーク値をとる周波数とを比較して得られた周波数の差分値を特徴量として抽出することを特徴とするものであったり、入力される周波数スペクトルからピーク値を検出し、検出されたピーク値をとる周波数と遅延手段によってある一定時間遅延させたピーク値をとる周波数とを比較し、ピーク値をとる周波数の時間軸方向の増減あるいはピーク値をとる周波数の変化の有無を2値の特徴量として抽出することを特徴

とするものであったり、入力される周波数スペクトルからピーク値を検出し、検出されたピーク値を異なる 2 つの帯域間で比較して得られた差分値を特徴量として抽出することを特徴とするものであったり、入力される周波数スペクトルからピーク値を検出し、検出されたピーク値を隣接する帯域間で比較し、ピーク値の周波数軸方向の増減を 2 値の特徴量として抽出することを特徴とするものであったり、入力される周波数スペクトルの実効値を計算し、算出された実効値を特徴量として抽出することを特徴とするものであったり、入力される周波数スペクトルの実効値を計算し、算出された実効値と遅延手段によって一定時間遅延させた実効値とを比較し、実効値の時間軸方向の増減や実効値の変化の有無を 2 値の特徴量として抽出することを特徴とするものであったり、入力される周波数スペクトルと遅延手段によって一定時間遅延させた周波数スペクトルとの相関値を計算し、算出された相関値を特徴量として抽出することを特徴とするものであったり

入力される周波数スペクトルと遅延手段によって一定時間遅延させた周波数スペクトルとの相関値を計算し、算出された相関値の符号を 2 値の特徴量として抽出することを特徴とするものであったり、入力される周波数スペクトルと遅延手段によって一定時間遅延させた周波数スペクトルとの相関値を計算し、算出された相関値と一定時間遅延させた相関値とを比較して得られた相関値の時間軸方向の増減を 2 値の特徴量として抽出することを特徴とするものである。

【 0 0 1 8 】

また、本発明の特徴量抽出装置は、音声信号をある一定の時区間ごとに切出す信号切出し手段と、前記信号切出し手段によって切出された信号を一定時間遅延させる信号遅延手段と、前記信号遅延手段から出力される信号と前記信号切出し手段によって切出された信号を入力して特徴量を抽出する抽出手段とから構成される。

【 0 0 1 9 】

また、前記抽出手段は、信号切出し手段の出力信号と信号遅延手段の出力信号との相関値を計算する相関演算手段を有し、前記相関演算手段によって算出される相関値を特徴量として抽出することを特徴とするものであったり、前記相関演

算手段によって算出される相関値の符号を 2 値の特徴量として抽出することを特徴とするものであったり、あるいは、相関演算手段によって算出された相関値を一定時間遅延させる相関値遅延手段を有し、前記相関演算手段から出力される相関値と前記相関値遅延手段から出力される相関値とを比較し、相関値の時間軸方向の増減を 2 値の特徴量として抽出することを特徴とするものである。

【 0 0 2 0 】

また、本発明の特徴量抽出装置は、音声信号から周波数スペクトルを求める周波数変換手段と、前記周波数変換手段によって求めた前記周波数スペクトルの包絡線を求めるための包絡線検出手段と、前記包絡線検出手段によって得られた前期周波数スペクトルの包絡線から特徴量を抽出する抽出手段とから構成される。

【 0 0 2 1 】

また、前記抽出手段は、入力される周波数スペクトルの包絡線から極大値あるいは極小値を検出し、検出された極大値あるいは極小値をとる周波数を特徴量として抽出することを特徴とするものであったり、あるいは、検出された極大値あるいは極小値をとる周波数を入力し、隣接する極値間の間隔を計算する間隔算出手段を有し、前記間隔算出手段によって得られる間隔を特徴量として抽出することを特徴とするものであったり、前記間隔検出手段によって得られた間隔に対して、前記極値検出手段によって得られた最も低い周波数と 2 番目に低い周波数との間隔などを基準値とした間隔比を特徴量として抽出することを特徴とするものである。

【 0 0 2 2 】

【発明の実施の形態】

以下、図面を参照して、本発明の実施の形態について説明する。尚、以下の実施の形態は、本発明を具体化した一例であって、本発明の技術的範囲を限定する性格のものではない。

【 0 0 2 3 】

(実施の形態 1)

図 1 から図 9 を用いて実施の形態 1 の特徴量抽出装置について説明する。

【 0 0 2 4 】

本実施の形態においては、入力された音声信号は、まず周波数変換手段 0 1 0 1 によって音声信号を周波数解析し、周波数スペクトルが求められる。ここで、周波数解析を行うに当たっては、高速フーリエ変換を用いる場合、音声信号から有限個のサンプル点を用いて演算を行うため、演算処理を行う前に音声信号からサンプル点数に相当する時間長の切出しを行う。周波数変換手段 0 1 0 1 にて求められた周波数スペクトルは、続いて、帯域分割手段 0 1 0 2 に入力され、特徴量を抽出するための複数の帯域に分割される。そして、帯域分割手段 0 1 0 2 によって分割された各帯域のスペクトルは抽出手段 0 1 0 3 に入力され、それぞれの帯域から特徴量の抽出が行われる。ここで、帯域分割手段 0 1 0 2 の動作について詳しい説明を行う。

【 0 0 2 5 】

帯域分割手段 0 1 0 2 にて行う分割の方法としては、周波数を線形目盛り上で等間隔に分割する方法が最も単純で効率的な方法である。また、音声信号の場合、音階の調和などの性質を考慮して周波数を対数目盛り上で等間隔に分割すると音楽的特徴の抽出に有効であることが考えられる。

【 0 0 2 6 】

また、帯域分割手段 0 1 0 2 では、図 2 に示すように任意の方法にて分割された帯域のうち、特定の帯域を選択して用いるようにしても良い。例えば、MP 3 のような圧縮技術を用いて符号化されている音声信号では、人間の聴覚特性上聞き取れないような高い周波数帯の情報を削除するといった操作がなされている。このような場合、情報が削除される高帯域の特徴量は音声指紋としての役割を果たし得ない。このように、特徴量が削除される可能性の高い帯域を除いて、替わりに分割された帯域の中から特徴量の抽出に有効なものを選択することによって、不必要な処理をなくして、少ない演算量にて音声信号の特定に有効な特徴量のみを抽出することができる。

【 0 0 2 7 】

また、帯域分割手段 0 1 0 2 では、図 3 に示すように分割された各帯域が互いに接するところや重なる領域がないように離散的に分割されても良い。例えば、図 2 に示すような境界線を含むような分割方法では、境界付近に特徴的な量が存

在する場合、外部からの雑音、擾乱、あるいは加工などによってスペクトルに歪みやずれが生じてしまうと隣接する帯域に対して影響を及ぼす可能性がある。したがって、帯域を離散的に設定することにより、周波数スペクトルの多少の変動を帯域間の領域（緩衝帯域と呼ぶ）の許容誤差の範囲内として吸収し、特徴量の抽出における強健性の向上を図ることが可能である。すなわち、外部からの要因によりスペクトルにずれが生じ、ピーク値のような特徴的な量が帯域から外れてしまう場合においても、前記緩衝帯域を設けておくことにより、ピーク値が隣接する帯域に含まれて特徴量に変化する事態を起こりにくくすることができる。

【 0 0 2 8 】

以上のようにして周波数スペクトルを複数の帯域に分割することは、1つのスペクトルの各帯域より特徴量を抽出することによって、帯域分割をしない場合より少ない演算量にて多くの特徴量を抽出できるため、その結果、多くの特徴量を用いてより精度の高い信号の検出に利用することが可能である。また、2つ以上に分割された帯域に対して、複数の帯域スペクトルを用いて処理を行うことにより新たな音声信号に関する特徴量を導出することも可能である。

【 0 0 2 9 】

なお、本実施の形態において、帯域分割の規則性は以上に示した方法に限定されるものではなく、あらゆる分割方法を含むものである。

【 0 0 3 0 】

続いて、抽出手段 0 1 0 3 の動作について詳しい説明を行う。帯域分割手段 0 1 0 2 によって分割された各帯域のスペクトルが抽出手段 0 1 0 3 に入力されると、抽出手段 0 1 0 3 では、図 4 に示すように、分割された帯域ごとにスペクトルの大きさが最も大きい値を示すものを探し出す処理が行われる。そして、分割された各帯域内でそれぞれ最もスペクトルの値が大きいところを検出すると、その値をとる周波数を特徴量として出力する。この方法によって、特徴量の検出が容易であり、かつ異なる音声信号との比較において十分な識別が可能な特徴量を抽出することができる。

【 0 0 3 1 】

また、抽出手段 0 1 0 3 では、図 5 に示すようにピーク検出手段 0 5 0 1 にて

検出されたスペクトルのピーク値をとる周波数を遅延手段 0 5 0 2 に入力して一定時間遅延させて、一定時間後、新たにピーク検出手段 0 5 0 1 にて検出されたピーク値をとる周波数と遅延手段 0 5 0 2 から出力される周波数とを比較手段 0 5 0 3 にて比較して特徴量を抽出するものでも良い。この場合、比較手段 0 5 0 3 では、まず、ピーク検出手段 0 5 0 1 から出力される周波数と遅延手段 0 5 0 2 から出力される周波数の差が計算される。すなわち、一定時間だけ離れたピーク値をとる周波数の時間変化量が求められる。比較手段 0 5 0 3 の出力としては、ピーク検出手段 0 5 0 1 の出力から遅延手段 0 5 0 2 の出力を引いた差分値をそのまま出力するものであったり、前記差分値の符号が正であれば 1、負であれば 0 となるような 2 値を出力するものであったり、差分値がある一定値を超えれば 1、そうでなければ 0 となるような 2 値を出力するものがある。差分値をそのまま出力するものは、すなわち、ピーク値をとる周波数の時間変化量の特徴量とするものである。差分値の符号を見て 2 値を出力するものは、すなわち、正ならば周波数が増加しており負ならば減少していることを意味しており、周波数の時間軸上の増減を特徴量とするものである。差分値が一定値を超えるかどうかで 2 値を出力するものは、すなわち、周波数のある閾値に対する変化の有無を意味するものであり、周波数が時間軸上で変化したか不変であるかを特徴量とするものである。この方法によって、特徴量の検出が容易であり、かつ異なる音声信号との比較において十分な識別が可能な特徴量を抽出することができる。また、時間差を特徴量とすることにより、音声信号の時間軸上の変化に対して耐性のある識別が可能である。

【 0 0 3 2 】

また、抽出手段 0 1 0 3 では、図 6 に示すように第 1 のピーク検出手段 0 6 0 1 にて検出された第 1 のピーク値と第 2 のピーク検出手段 0 6 0 2 にて検出された第 2 のピーク値とを比較手段 0 6 0 3 にて比較して特徴量を抽出するものでも良い。この場合、比較手段 0 6 0 3 では、隣接する 2 つの帯域間におけるピーク値の差分値が計算される。前記差分値は、すなわち、帯域間におけるピーク値の変化量を意味するものである。比較手段 0 6 0 3 の出力としては、差分値をそのまま出力するものであったり、前記差分値の符号が正であれば 1、負であれば 0

となるような2値を出力するものがある。差分値をそのまま出力するものは、各帯域間のピーク値の変化量を特徴量とするものである。差分値の符号をみて2値を出力するものは、すなわち、スペクトルの周波数軸方向に対する帯域間のピーク値の増減を特徴量とするものである。この方法によって、特徴量の検出が容易であり、かつ異なる音声信号との比較において十分な識別が可能な特徴量を抽出することができる。また、帯域間のスペクトル差を特徴量とすることにより、音声信号の周波数軸上の変化に対して耐性のある識別が可能である。

【 0 0 3 3 】

なお、本実施の形態において、ピーク値の差分値を計算する2つの帯域は、以上に示した隣接する2つの帯域に限定するものではなく、分割された複数の帯域の中から選ばれた任意の2つの帯域に対する場合を含むものである。

【 0 0 3 4 】

また、抽出手段0103では、分割された帯域ごとに各帯域中のスペクトルの実効値、すなわち、スペクトルの二乗平均平方根を計算し、特徴量として出力するものでも良い。この方法によって、特徴量の検出が容易であり、かつ異なる音声信号との比較において十分な識別が可能な特徴量を抽出することができる。

【 0 0 3 5 】

また、抽出手段0103では、図7に示すように実効値計算手段0701にて検出された実効値を遅延手段0702に入力して一定時間遅延させて、一定時間後、新たに実効値計算手段0701にて検出された実効値と遅延手段0702から出力される実効値とを比較手段0703にて比較して特徴量を抽出するものでも良い。この場合、比較手段0703では、まず、実効値計算手段0701から出力される実効値と遅延手段0702から出力される実効値の差が計算される。すなわち、一定時間だけ離れた実効値の時間変化量が求められる。比較手段0703の出力としては、実効値計算手段0701の出力から遅延手段0702の出力を引いた差分値をそのまま出力するものであったり、差分値の符号が正であれば1、負であれば0となるような2値を出力するものであったり、差分値がある一定値を超えれば1、そうでなければ0となるような2値を出力するものがある。差分値をそのまま出力するものは、すなわち、実効値の時間変化量を特徴量と

するものである。差分値の符号を見て 2 値を出力するものは、すなわち、正ならば実効値が増加しており負ならば減少していることを意味するため実効値の時間軸上の増減を特徴量とするものである。差分値が一定値を超えるかどうかで 2 値を出力するものは、すなわち、実効値のある閾値に対する変化の有無を意味するものであり、実効値が時間軸上で変化したか不変であるかを特徴量とするものである。この方法によって、特徴量の検出が容易であり、かつ異なる音声信号との比較において十分な識別が可能な特徴量を抽出することができる。また、時間差を特徴量とすることにより、音声信号の時間軸上の変化に対して耐性のある識別が可能である。

【 0 0 3 6 】

また、抽出手段 0 1 0 3 では、図 8 に示すように一定時間だけ離れた 2 つの周波数スペクトル間の相関から特徴量を抽出するものでも良い。すなわち、抽出手段 0 1 0 3 に入力される周波数スペクトルをスペクトル遅延手段 0 8 0 1 にて一定時間遅延させて、一定時間後、新たに入力される周波数スペクトルとスペクトル遅延手段 0 8 0 1 から出力される周波数スペクトルとの相関値を相関演算手段 0 8 0 2 において計算し、特徴量を抽出する。この場合、相関演算手段 0 8 0 2 の出力としては、相関値をそのまま特徴量として出力するものであったり、相関値の符号が正であれば 1、負であれば 0 となるような 2 値を出力するものがある。この方法によって、異なる音声信号との比較において十分な識別が可能な特徴量を抽出することができ、また、音声信号内の時間差に関する特徴量であり、音声信号の時間軸上の変化に対して耐性のある識別が可能である。

【 0 0 3 7 】

また、抽出手段 0 1 0 3 では、図 9 に示すように一定時間だけ離れた 2 つの周波数スペクトル間の相関値を求め、相関値の時間差を特徴量とするものでも良い。すなわち、まず、抽出手段 0 1 0 3 に入力される周波数スペクトルをスペクトル遅延手段 0 9 0 1 にて一定時間遅延させて、一定時間後、新たに入力される周波数スペクトルとスペクトル遅延手段 0 9 0 1 から出力される周波数スペクトルとの相関値を相関演算手段 0 9 0 2 において計算する。さらに、算出された相関値を相関値遅延手段 0 9 0 3 にて一定時間遅延させて、一定時間後、相関演算手

段 0 9 0 2 より出力される相関値と相関値遅延手段 0 9 0 3 から出力される相関値とを比較手段 0 9 0 4 にて比較し、特徴量を抽出する。この場合、比較手段 0 9 0 4 では、相関演算手段 0 9 0 2 から出力される相関値と相関値遅延手段 0 9 0 3 から出力される相関値と差が計算される。すなわち、一定時間だけ離れた相関値の時間変化量が求められる。比較手段 0 9 0 4 の出力としては、相関演算手段 0 9 0 2 の出力から相関値遅延手段 0 9 0 3 の出力を引いた差分値の符号が正であれば 1、負であれば 0 となるような 2 値を出力するものである。差分値の符号を見て 2 値を出力するものは、すなわち、正ならば相関値が増加しており負ならば減少していることを意味するため相関値の時間軸上の増減を特徴量とするものである。この方法によって、異なる音声信号との比較において十分な識別が可能な特徴量を抽出することができ、また、時間差を特徴量とすることにより、音声信号の時間軸上の変化に対して耐性のある識別が可能である。

【 0 0 3 8 】

以上のようにして抽出された特徴量は、前記音声信号を他の異なる信号と識別することができるものであり、特徴量を基にして未知の音声信号を特定することが可能である。

【 0 0 3 9 】

(実施の形態 2)

図 1 0 から図 1 2 を用いて実施の形態 2 の特徴量抽出装置について説明する。

【 0 0 4 0 】

本実施の形態においては、入力される音声信号に対して音声信号内の時間が異なる部分間での相関値を計算し、特徴量とするものである。まず、入力された音声信号は、信号切出し手段 1 0 0 1 によって任意の設定された時間長にて切出しが行われる。続いて、信号切出し手段 1 0 0 1 にて切出された信号は、一旦信号遅延手段 1 0 0 2 においてある一定時間遅延される。信号遅延手段 1 0 0 2 によって一定時間だけ遅延された信号は、新たに信号切出し手段によって切出された信号とともに抽出手段 1 0 0 3 に入力され、特徴量が抽出される。

【 0 0 4 1 】

ここで、抽出手段 1 0 0 3 の動作について詳しく説明する。

【 0 0 4 2 】

抽出手段 1 0 0 3 では、2 つの離れた時間の信号間で相関値の計算が行われる。例えば、図 1 1 に示すように、信号切出し手段 1 0 0 1 において音声信号が時間軸上で一定の時間長（フレーム）ごとに切出されると、第 1 のフレーム信号と第 3 のフレーム信号のように 2 フレーム時間だけ離れた信号が抽出手段 1 0 0 3 に入力され、入力された 2 つの信号から相関値が計算され、算出された相関値が特徴量として抽出される。

【 0 0 4 3 】

なお、本実施の形態では相関値を計算する 2 つのフレーム信号の間隔は 2 フレーム時間に限定されるものではなく、任意の時間だけ離れた 2 つの信号との間で相関値が計算される場合を含むものとする。

【 0 0 4 4 】

抽出手段 1 0 0 3 の出力としては、相関値をそのまま出力するものであったり、相関値の符号が正であれば 1、負であれば 0 となるような 2 値を出力するものがある。

【 0 0 4 5 】

なお、フレーム信号の切出すタイミングは図 1 1 に示すようなフレーム時間ごとに限定されるものではなく、あらゆるタイミングを含むものである。この方法によって、特徴量の検出が容易であり、かつ異なる音声信号との比較において十分な識別が可能な特徴量を抽出することができる。また、音声信号内の時間差に関する特徴量であり、音声信号の時間軸上の変化に対して耐性のある識別が可能である。

【 0 0 4 6 】

また、抽出手段 1 0 0 3 では、図 1 2 に示すように相関演算手段 1 2 0 1 にて算出された相関値を相関値遅延手段 1 2 0 2 によって一定時間遅延させて、一定時間後、新たに相関演算手段 1 2 0 1 によって算出された相関値と相関値遅延手段 1 2 0 2 に保持しておいた相関値とを比較手段 1 2 0 3 にて比較して特徴量を抽出するものでも良い。比較手段 1 2 0 3 では、相関演算手段 1 2 0 1 から出力される相関値と相関値遅延手段 1 2 0 2 から出力される相関値との差が計算され

る。すなわち、一定時間はなれた相関値の時間変化量が求められる。比較手段 1 2 0 3 の出力としては、相関演算手段 1 2 0 1 から出力される相関値から相関値遅延手段 1 2 0 2 から出力される相関値を引いた差分値の符号が正であれば 1、負であれば 0 となるような 2 値を出力するものである。差分値の符号を見て 2 値を出力するものは、すなわち、正ならば相関値が増加しており負ならば減少していることを意味するため相関値の時間軸上の増減を特徴量とするものである。この方法によって、特徴量の検出が容易であり、かつ異なる音声信号との比較において十分な識別が可能な特徴量を抽出することができる。また、時間差を特徴量とすることにより、音声信号の時間軸上の変化に対して耐性のある識別が可能である。

【 0 0 4 7 】

(実施の形態 3)

図 1 3 および図 1 4 を用いて実施の形態 3 の特徴量抽出装置について説明する。本実施の形態においては、入力された音声信号は、まず周波数変換手段 1 3 0 1 によって前記音声信号を周波数解析し、周波数スペクトルが求められる。周波数変換手段 1 3 0 1 に関しては第 1 の実施の形態にて説明したものと同様であるため、説明を省略する。周波数変換手段 1 3 0 1 によって得られた周波数スペクトルは、続いて包絡線検出手段 1 3 0 2 に入力され、周波数スペクトルの包絡線が検出される。周波数スペクトルの包絡線を得ることによって、音声信号の周波数領域のゆるやかな変動を捉えることができる。続いて、抽出手段 1 3 0 3 において、周波数スペクトルの包絡線から極値を検出し、検出された極値をとる周波数が出力される。

【 0 0 4 8 】

図 1 4 は、抽出手段 1 3 0 3 による検出の様子を示している。図 1 4 の左の図は周波数スペクトルの包絡線から極大値を検出し、その値をとる周波数を出力して特徴量とするものであり、図 1 4 の右の図は周波数スペクトルの包絡線から極小値を検出し、その値をとる周波数を出力して特徴量とするものである。

【 0 0 4 9 】

本実施の形態では、スペクトルの包絡線を求めることにより、周波数領域のゆ

るやかな変化を捉えやすくすることができ、したがって、極大値あるいは極小値のような周波数領域の特徴量を容易に検出することができる。

【 0 0 5 0 】

（実施の形態 4）

図 1 5 および図 1 6 を用いて実施の形態 4 の特徴量抽出装置について説明する。図 1 5 は実施の形態 3 を説明した図 1 3 に対して、抽出手段 1 3 0 3 において極値検出手段 1 5 0 1 にて検出された極値をとる周波数を間隔算出手段 1 5 0 2 に入力して、ある値を基準として極値をとる周波数の間隔比を求め、特徴量として抽出するものである。

【 0 0 5 1 】

図 1 6 は間隔算出手段 1 5 0 2 による検出の様子を示している。間隔算出手段 1 5 0 2 では、極値検出手段 1 5 0 1 にて検出された複数の周波数が入力されると、隣接する各周波数間の差が計算される。算出された周波数の差はその値をそのまま特徴量とすることもできるが、ある基準値に基づいた相対的な比を特徴量とすることができる。間隔算出手段 1 5 0 2 において間隔比を求める基準値としては、極値検出手段 1 5 0 1 にて検出された極値の中で最も低い周波数や、間隔算出手段 1 5 0 2 によって得られた複数の差の中から極値検出手段 1 5 0 1 によって得られた最も低い周波数と 2 番目に低い周波数との差などを用いる。

【 0 0 5 2 】

なお、本発明において、間隔比を求める基準値は、以上に示した検出された極値の中で最も低い周波数や、複数の差の中から極値検出手段によって得られた最も低い周波数と 2 番目に低い周波数との差に限定するものではなく、あらゆる値を基準値とする場合を含むものである。

【 0 0 5 3 】

このような極値をとる周波数間の間隔の比を特徴量とすることは、音声信号のテンポが変更されるなどによって周波数スペクトルが伸縮する場合に対して非常に強健な特徴量を得ることになる。

【 0 0 5 4 】

【発明の効果】

以上のように、本発明の特徴量抽出装置を用いることにより、音声信号から固有の特徴量を抽出することができ、その抽出された特徴量は、さまざまな加工や外部雑音に対してロバスト性の高い特徴量である。

【 0 0 5 5 】

また、本発明における特徴量の抽出手段は比較的容易に実現することができ、元の音声信号と比べて小さいデータ量にて表現される特徴量を抽出することができる。

【図面の簡単な説明】

【図 1】

本発明の実施の形態 1 の特徴量抽出装置を示すブロック図

【図 2】

本発明の実施の形態 1 の帯域分割手段の概略を示す図

【図 3】

本発明の実施の形態 1 の帯域分割手段の概略を示す図

【図 4】

本発明の実施の形態 1 の抽出手段の概略を示す図

【図 5】

本発明の実施の形態 1 の抽出手段の詳細を示すブロック図

【図 6】

本発明の実施の形態 1 の抽出手段の詳細を示すブロック図

【図 7】

本発明の実施の形態 1 の抽出手段の詳細を示すブロック図

【図 8】

本発明の実施の形態 1 の抽出手段の詳細を示すブロック図

【図 9】

本発明の実施の形態 1 の抽出手段の詳細を示すブロック図

【図 1 0】

本発明の実施の形態 2 の特徴量抽出装置を示すブロック図

【図 1 1】

本発明の実施の形態 2 の抽出手段の概略を示す図

【図 1 2】

本発明の実施の形態 2 の抽出手段の詳細を示すブロック図

【図 1 3】

本発明の実施の形態 3 の特徴量抽出装置を示すブロック図

【図 1 4】

本発明の実施の形態 3 の抽出手段の概略を示す図

【図 1 5】

本発明の実施の形態 4 に係る抽出手段の詳細を示すブロック図

【図 1 6】

本発明の実施の形態 4 に係る抽出手段の概略を示す図

【図 1 7】

従来の音声指紋技術における音声指紋情報を蓄積する過程を説明するブロック

図

【図 1 8】

従来の音声指紋技術における音声指紋情報を検出する過程を説明するブロック

図

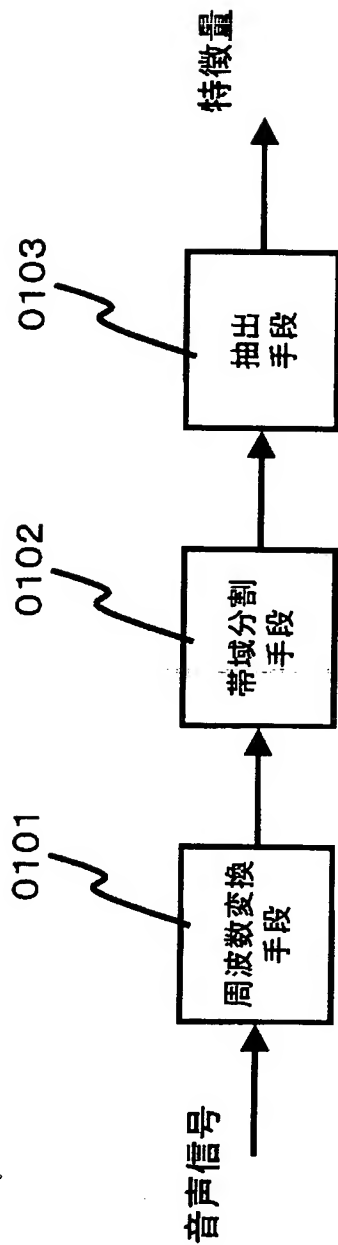
【符号の説明】

- 0 1 0 1 周波数変換手段
- 0 1 0 2 帯域分割手段
- 0 1 0 3 抽出手段
- 0 5 0 1 ピーク検出手段
- 0 5 0 2 遅延手段
- 0 5 0 3 比較手段
- 0 6 0 1 第 1 のピーク検出手段
- 0 6 0 2 第 2 のピーク検出手段
- 0 6 0 3 比較手段
- 0 7 0 1 実効値計算手段
- 0 7 0 2 遅延手段

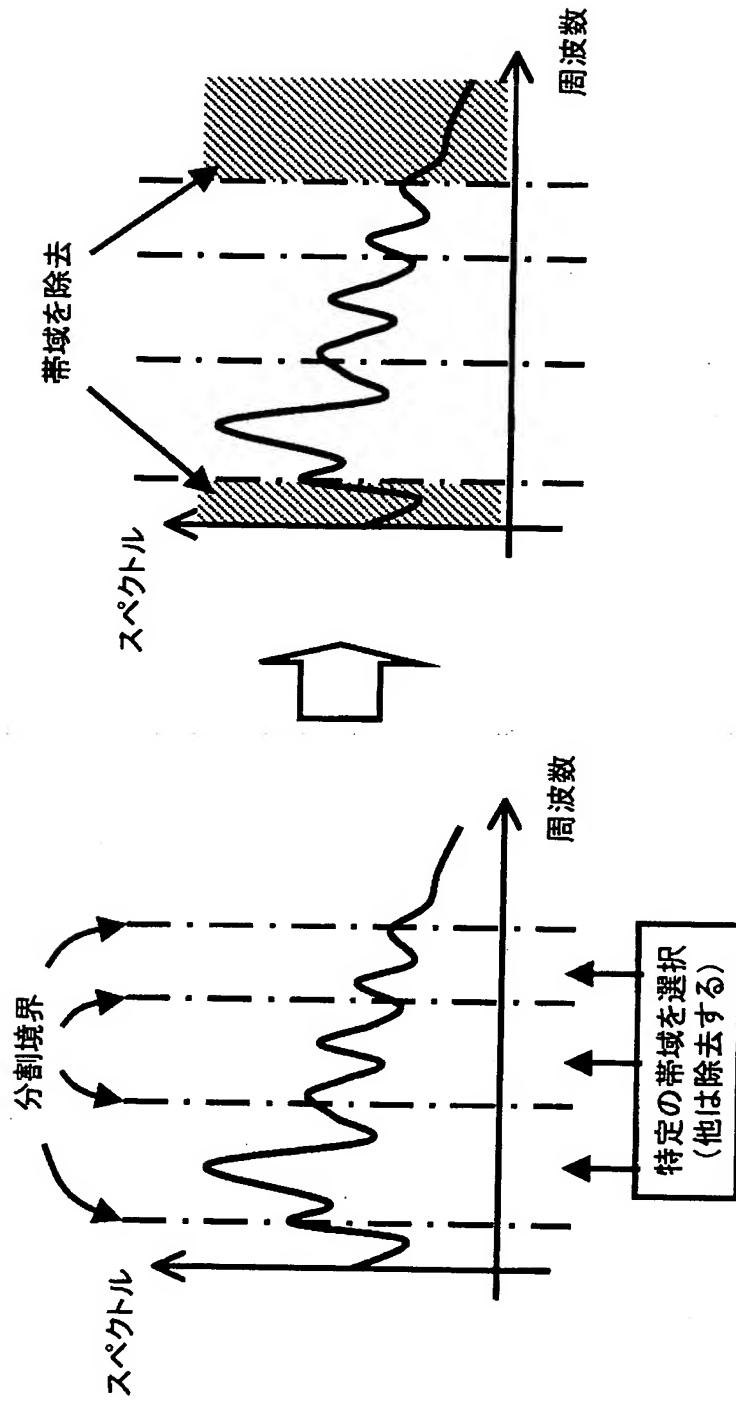
- 0 7 0 3 比較手段
- 0 8 0 1 スペクトル遅延手段
- 0 8 0 2 相関演算手段
- 0 9 0 1 スペクトル遅延手段
- 0 9 0 2 相関演算手段
- 0 9 0 3 相関値遅延手段
- 0 9 0 4 比較手段
- 1 0 0 1 信号切出し手段
- 1 0 0 2 信号遅延手段
- 1 0 0 3 抽出手段
- 1 2 0 1 相関演算手段
- 1 2 0 2 相関値遅延手段
- 1 2 0 3 比較手段
- 1 3 0 1 周波数変換手段
- 1 3 0 2 包絡線検出手段
- 1 3 0 3 抽出手段
- 1 5 0 1 極値検出手段
- 1 5 0 2 間隔算出手段
- 1 7 0 1 特徴量抽出装置
- 1 7 0 2 蓄積手段
- 1 8 0 1 特徴量抽出装置
- 1 8 0 2 蓄積手段
- 1 8 0 3 指紋比較手段

【書類名】 図面

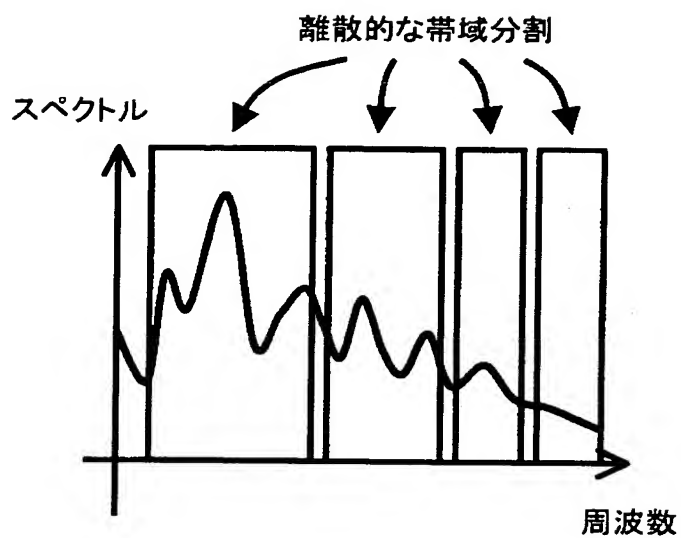
【図1】



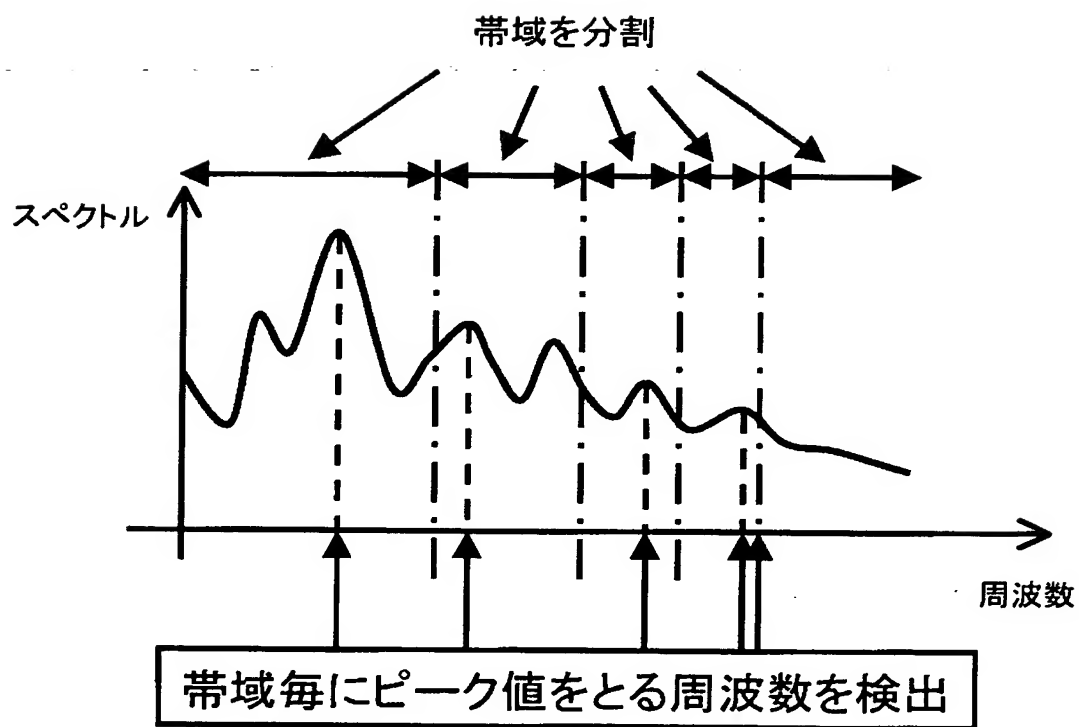
【図 2】



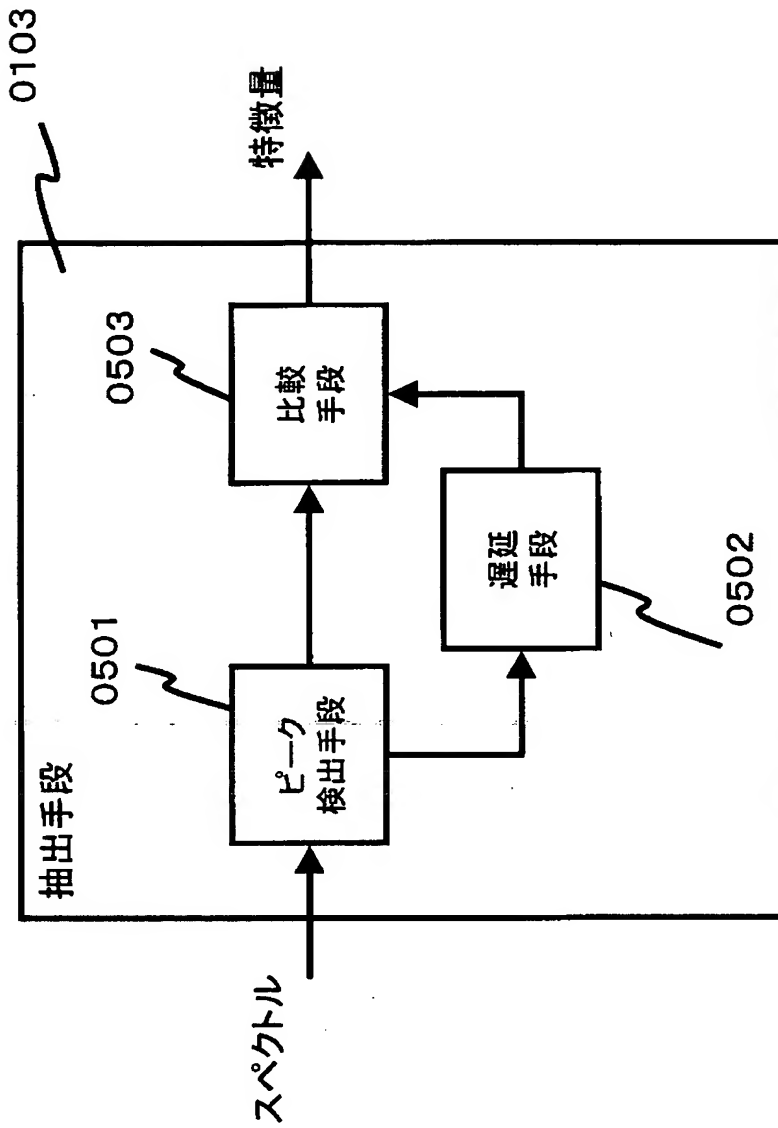
【図 3】



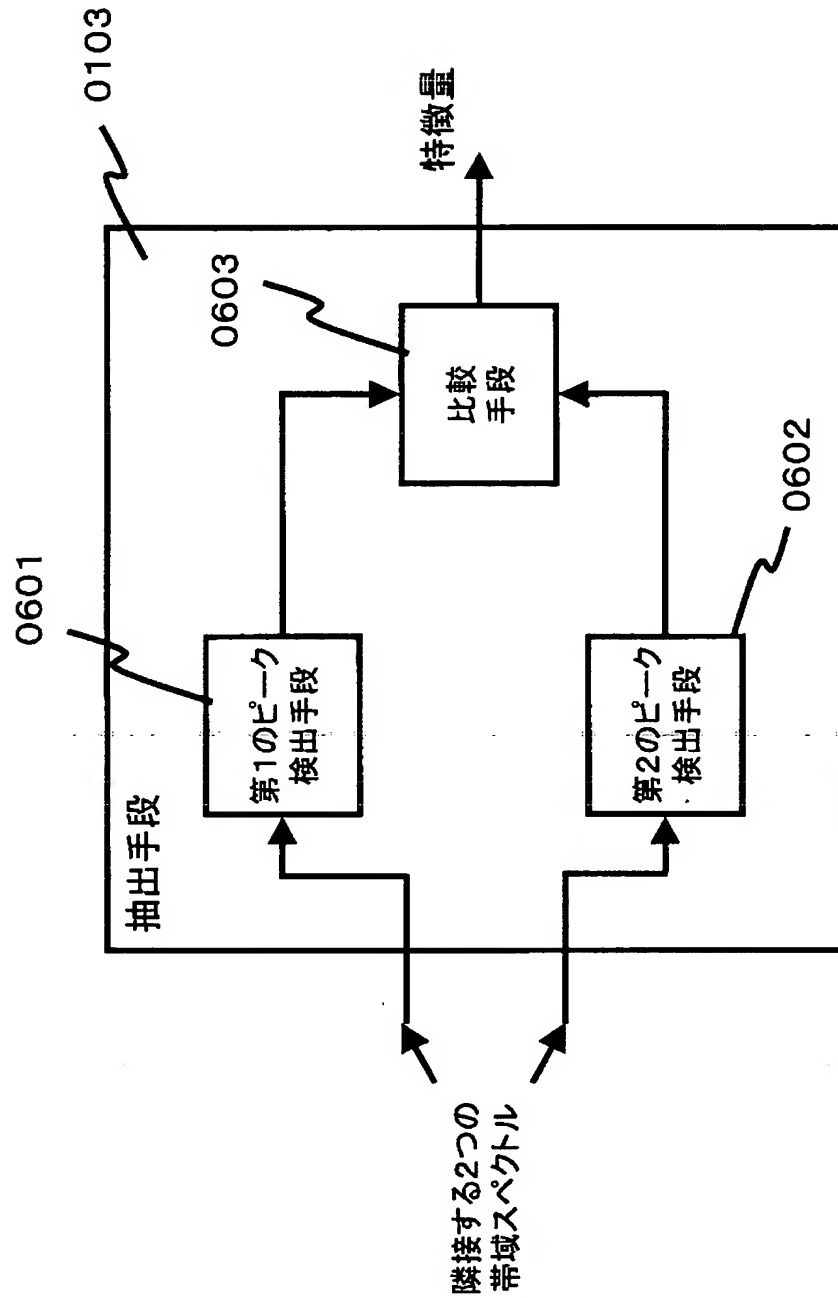
【図 4】



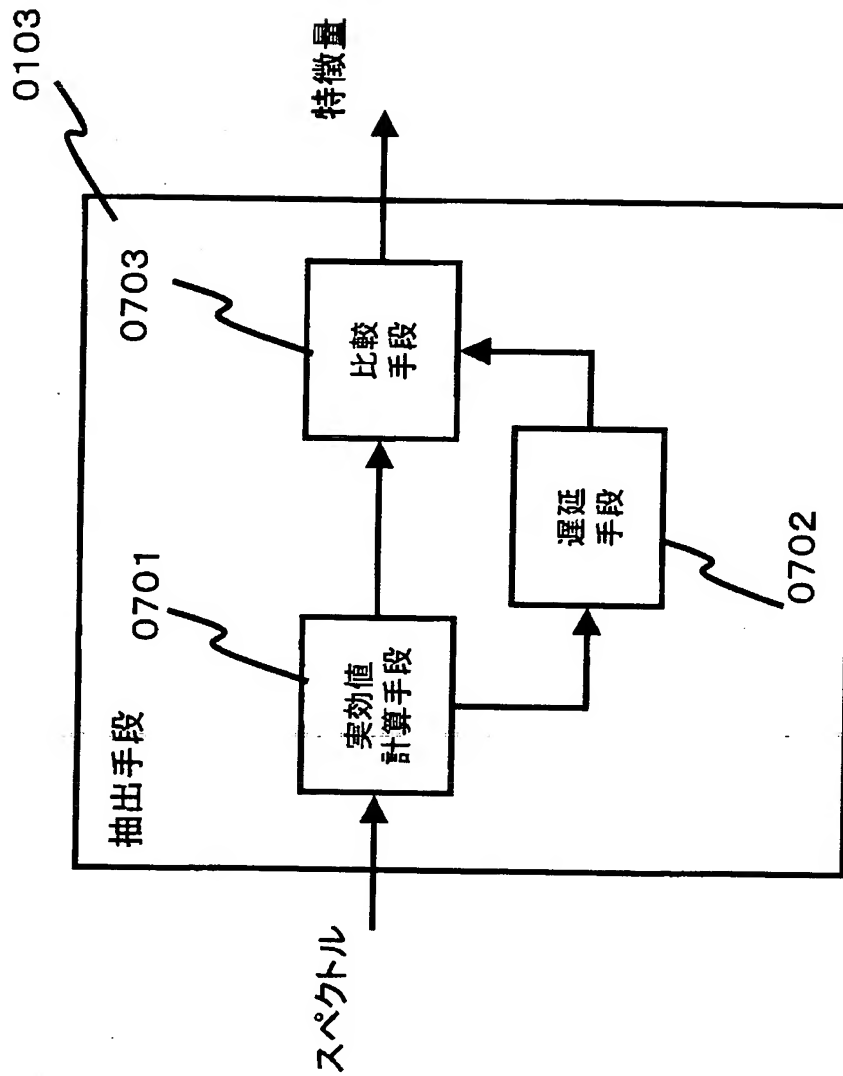
【図 5】



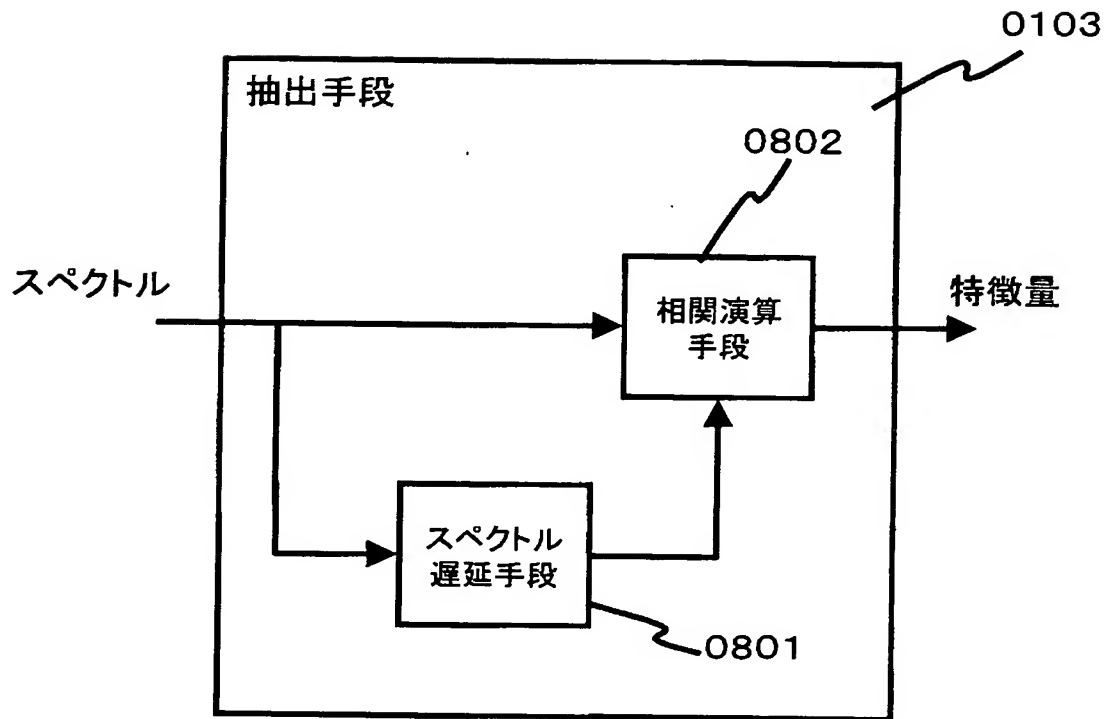
【図 6】



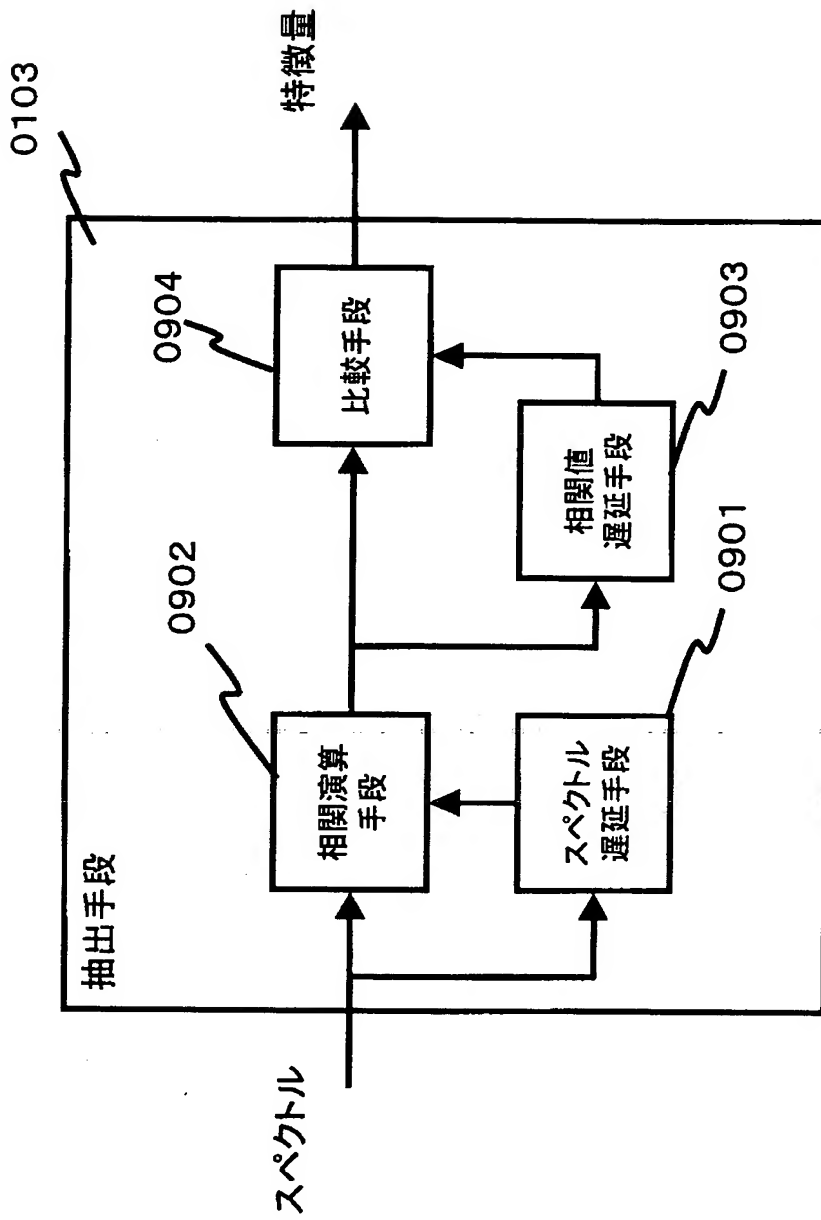
【図7】



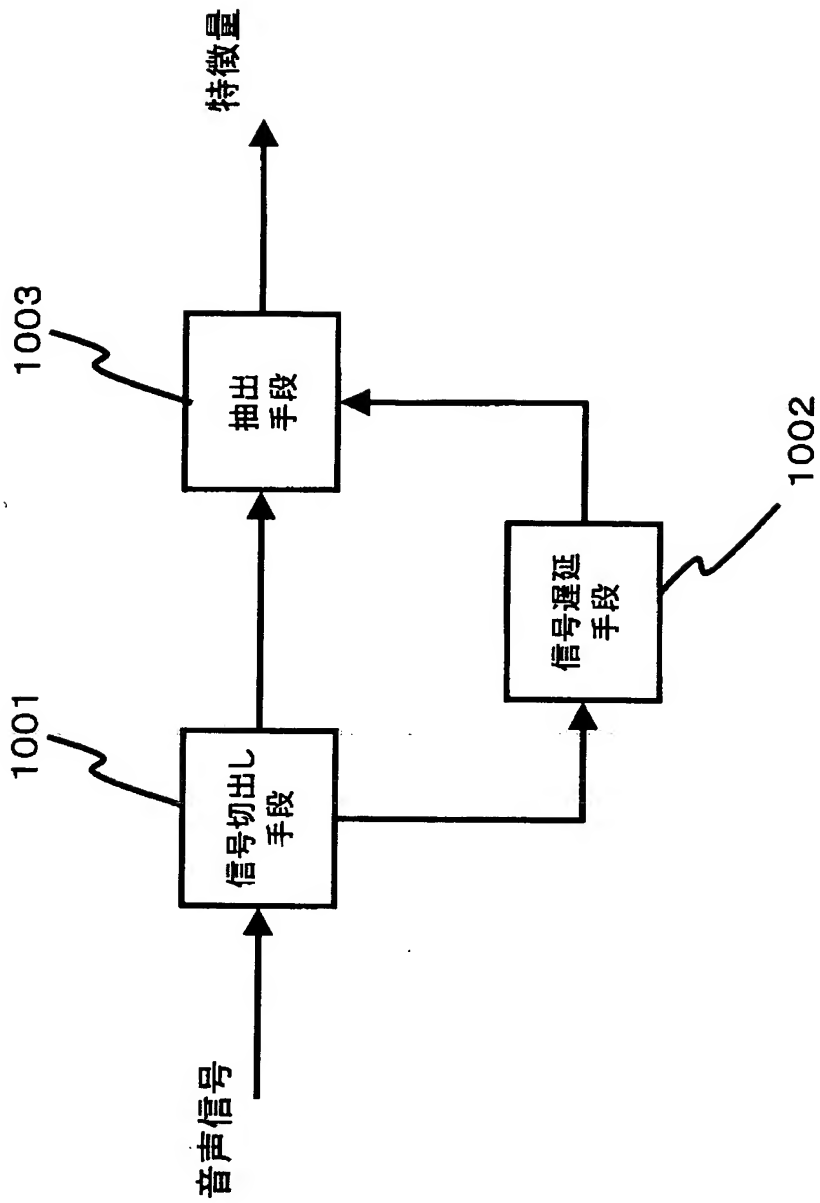
【図 8】



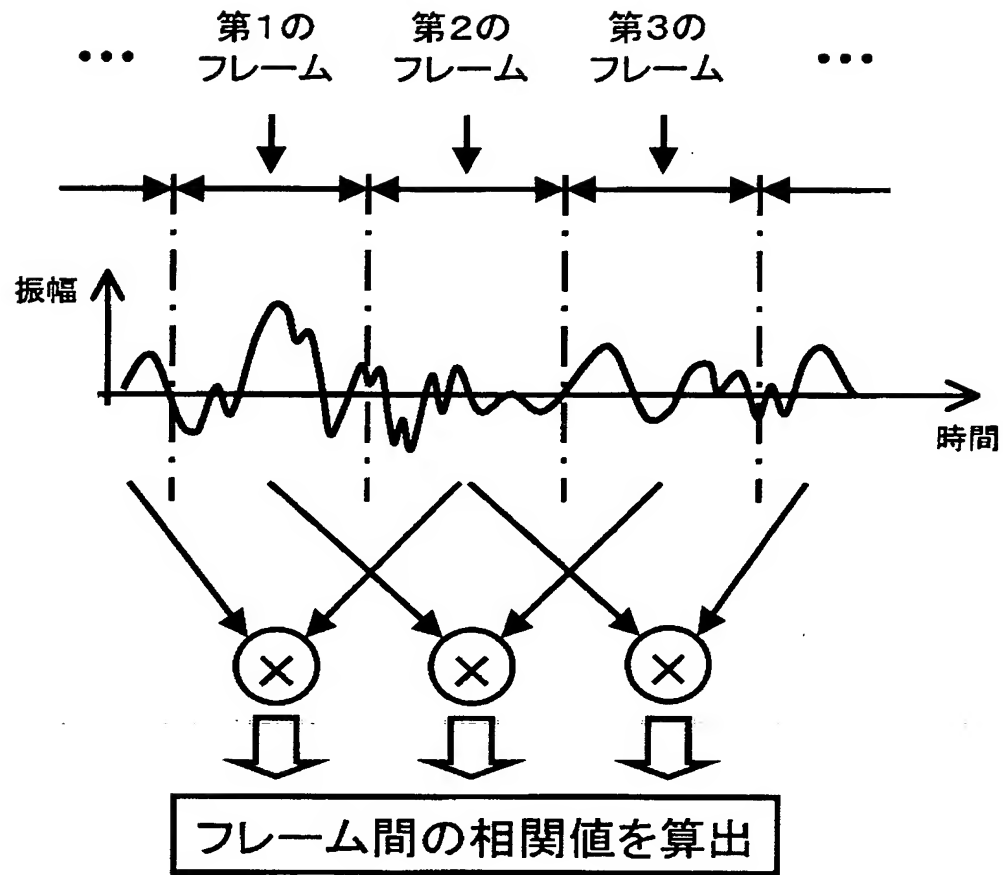
【図 9】



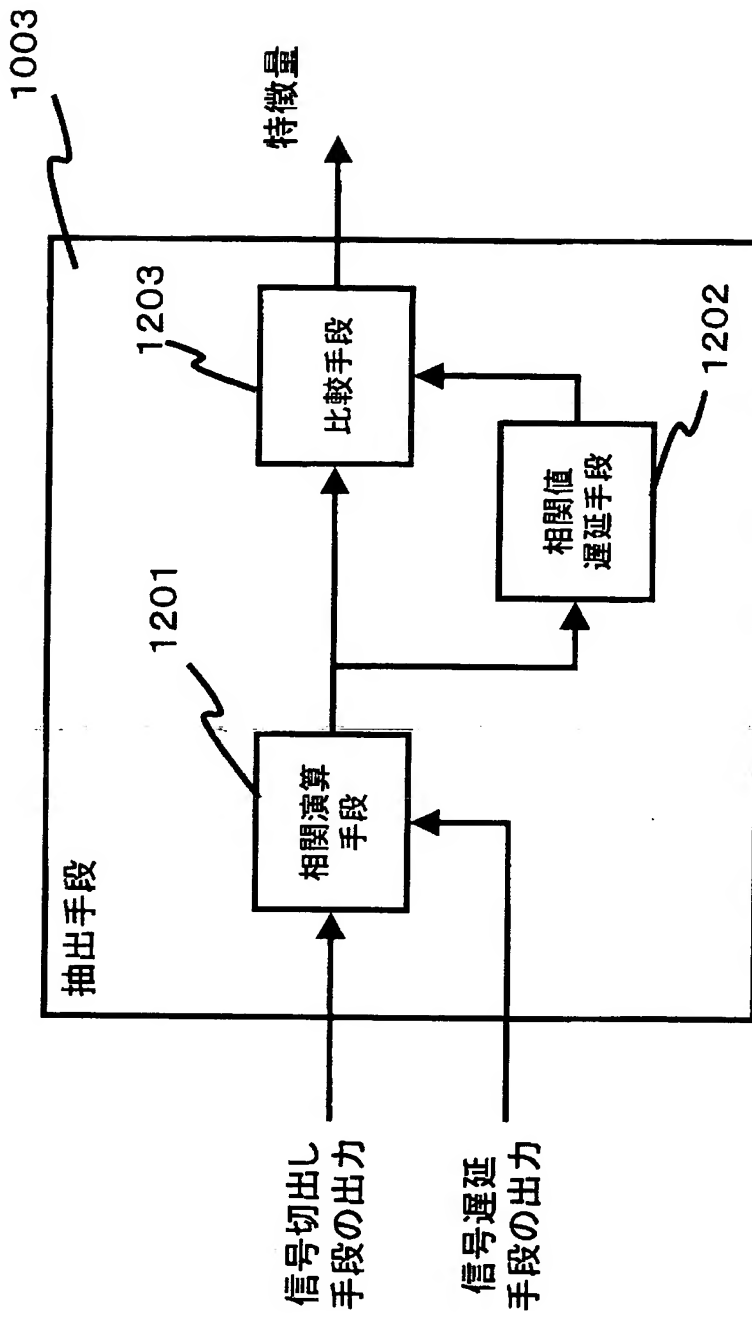
【図 1 0】



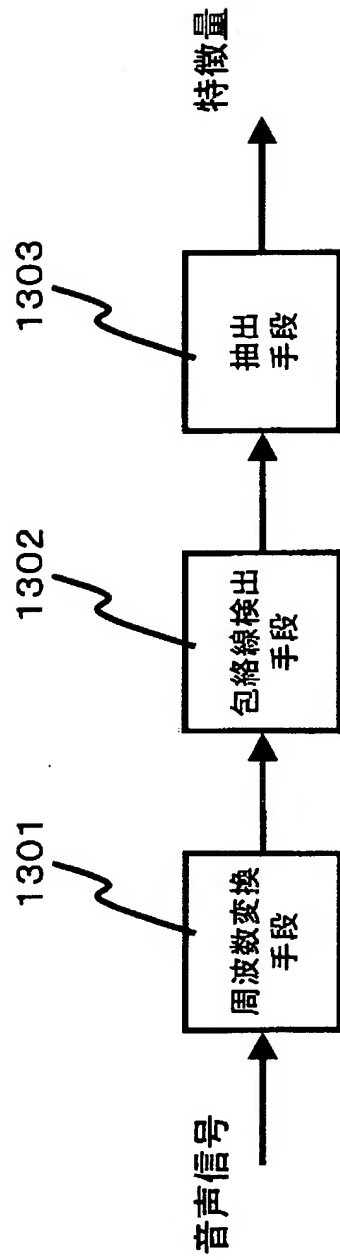
【図 1 1】



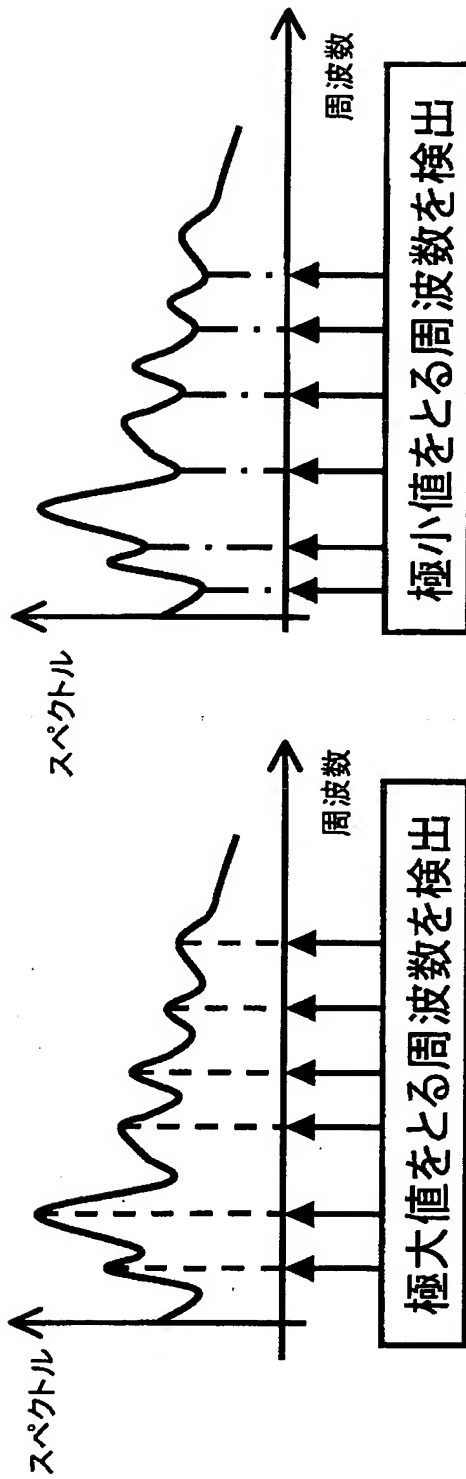
【図 12】



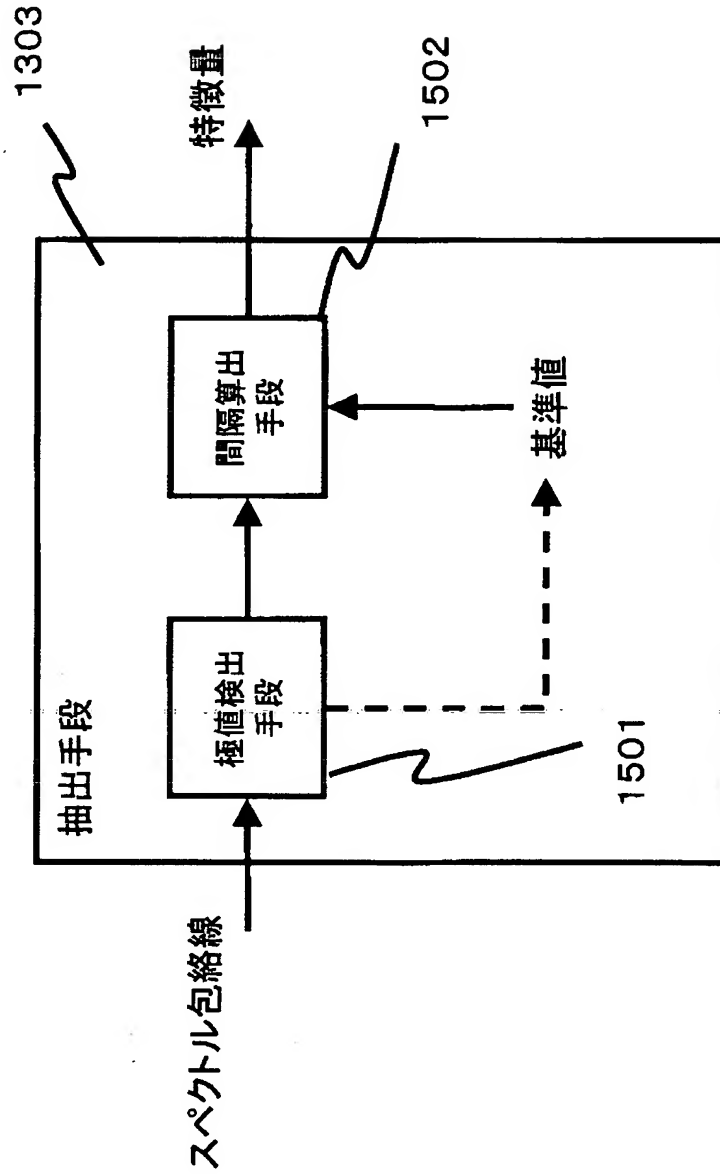
【図 1 3】



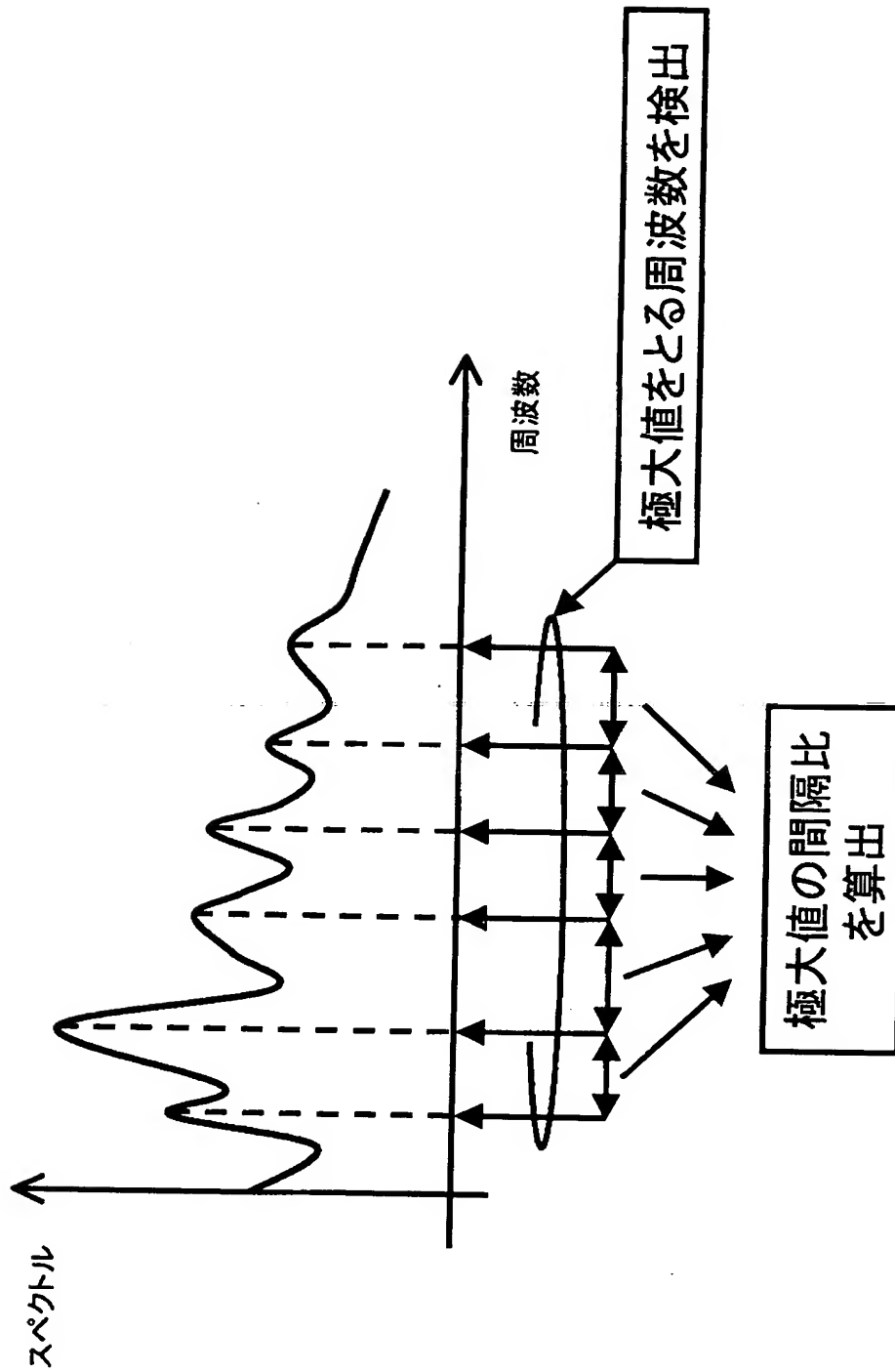
【図 14】



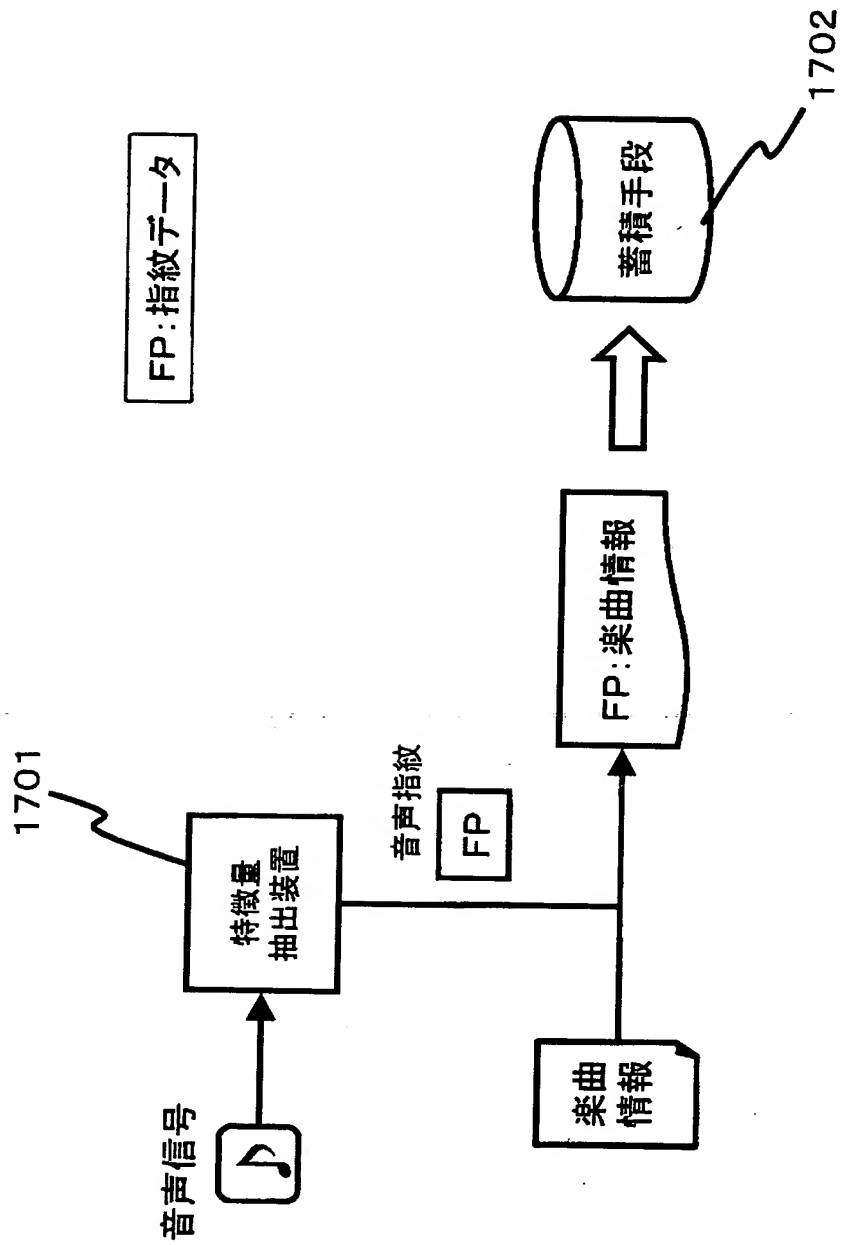
【図 1 5】



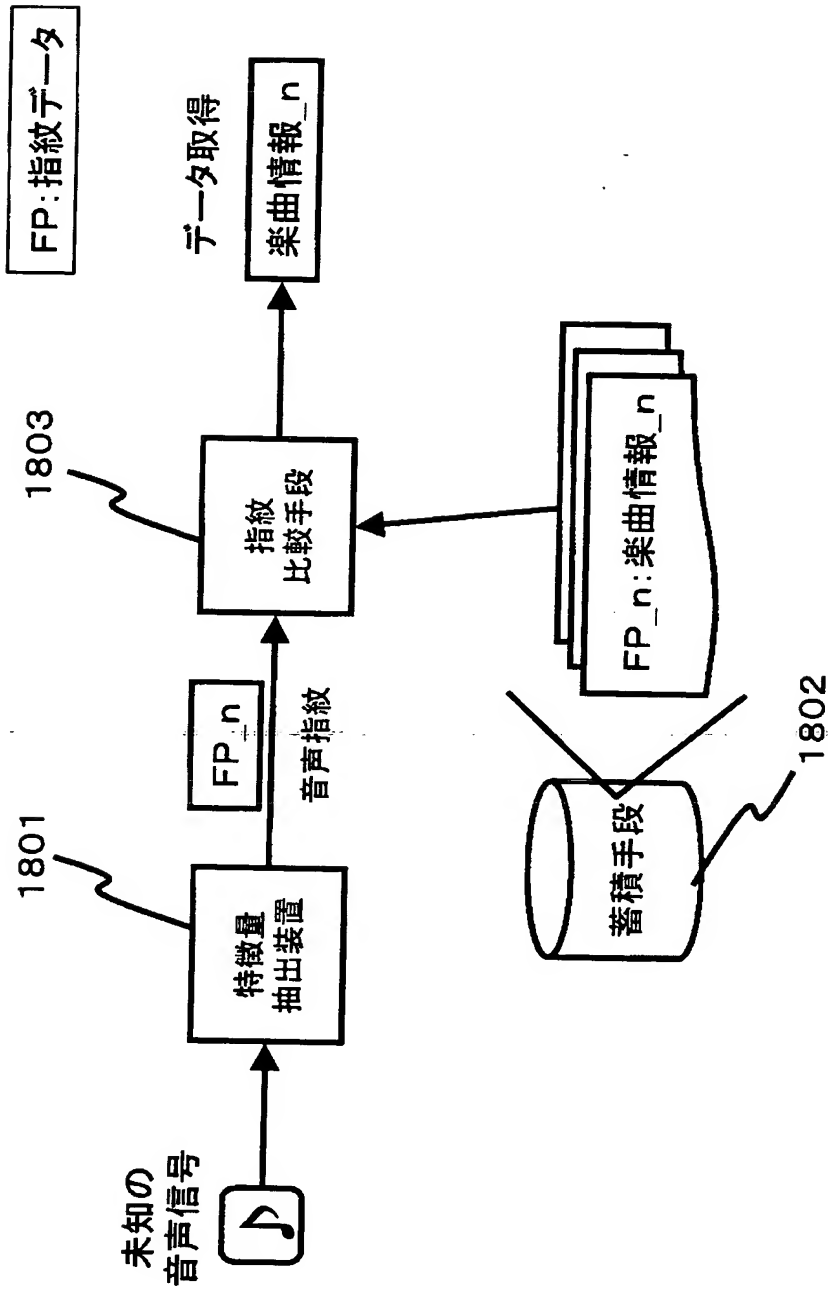
【図 1 6】



【図17】



【図 18】



【書類名】 要約書

【要約】

【課題】 音声指紋は、音声信号に多少の加工が加えられた信号についても加工前の元の音声信号と同等であることを検出でき、また、他の音声信号の指紋と比較したとき、明らかに異なる音声信号であることを識別することができる特徴量でなければならない。

【解決手段】 入力信号から周波数スペクトルを求める周波数変換手段と、前記周波数変換手段によって求めた前記周波数スペクトルを1つ以上の帯域に分割する帯域分割手段と、前記分割手段によって分割された帯域の各々から特徴量を抽出する抽出手段と、を具備することを特徴とする。

【選択図】 図1

出 願 人 履 歴 情 報

識別番号 [000005821]

1. 変更年月日	1990年 8月28日
[変更理由]	新規登録
住 所	大阪府門真市大字門真1006番地
氏 名	松下電器産業株式会社